

NAACL HLT 2019

**The 2019 Conference of the  
North American Chapter of the  
Association for Computational Linguistics:  
Human Language Technologies**

**Proceedings of the Demonstrations Session**

June 2 - June 7, 2019  
Minneapolis, MN

©2019 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)  
209 N. Eighth Street  
Stroudsburg, PA 18360  
USA  
Tel: +1-570-476-8006  
Fax: +1-570-476-0860  
[acl@aclweb.org](mailto:acl@aclweb.org)

ISBN 978-1-950737-16-1 (Demos)

## **Introduction**

Welcome to the system demonstrations track at NAACL-HLT 2019. We received 63 submissions in total, 24 of which were accepted.

We thank all the authors who submitted their work, and our program committee members for their hard work in helping us select an excellent set of demos.



**Organizers:**

Waleed Ammar, Allen Institute for Artificial Intelligence  
Annie Louis, University of Edinburgh  
Nasrin Mostafazadeh, Elemental Cognition

**Program Committee:**

Puneet Agarwal, TCS Research  
Tamer Alkhouli, RWTH Aachen University  
Omar Alonso, Microsoft  
Pepa Atanasova, University of Copenhagen  
Tyler Baldwin, IBM Research - Almaden  
Guntis Barzdins, University of Latvia  
Iz Beltagy, Allen Institute for Artificial Intelligence  
Akash Bharadwaj, Elemental Cognition  
Mikhail Burtsev, Moscow Institute of Physics and Technology  
Miles Crawford, Allen Institute for Artificial Intelligence  
Montse Cuadros, Vicomtech  
Thierry Declerck, DFKI GmbH  
Tim Dettmers, University of Washington  
Lipika Dey, TCS  
Pradeep Dogga, UCLA  
Eduard Dragut, Temple University  
Roald Eiselen, South African Centre for Digital Language Resources (SADiLaR)  
Emad Elwany, Allen Institute for Artificial Intelligence  
Debasis Ganguly, IBM Research Lab, Dublin  
Ulrich Germann, University of Edinburgh  
Martin Gleize, IBM Research Ireland  
Ajda Gokcen, The University of Washington  
Maarten van Gompel, Radboud University Nijmegen  
Frank Grimm, Semantic Computing Group, CITEC, Bielefeld University  
Joel Grus, Allen Institute for Artificial Intelligence  
Suchin Gururangan, Allen Institute for AI  
Francisco Guzmán, Facebook  
Jingguang Han, Accenture Labs  
Jer Hayes, Accenture Labs  
Ari Holtzman, University of Washington  
Charles Jochim, IBM Research  
Sujeong Kim, SRI International  
Roman Klinger, University of Stuttgart  
Sunil Kumar Koppurapu, TCS Research and Innovation - Mumbai  
Brigitte Krenn, Austrian Research Institute for Artificial Intelligence  
Vishwajeet Kumar, IITB-Monash Research Academy  
Ni Lao, mosaix.ai  
Chen Li, Tencent  
Zuchao Li, Shanghai Jiao Tong University  
Yiyun Liang, University of Waterloo  
Ying Lin, Rensselaer Polytechnic Institute  
Wei Liu, The University of Western Australia  
Nicholas Lourie, Allen Institute for Artificial Intelligence

Di Lu, Rensselaer Polytechnic Institute  
Ayush Maheshwari, IIT Bombay  
Yuji Matsumoto, Nara Institute of Science and Technology  
Jonathan May, USC Information Sciences Institute  
James Mayfield, Johns Hopkins University  
Marie-Jean Meurs, Université du Québec à Montréal  
Ivan Vladimir Meza Ruiz, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS-UNAM)  
Pasquale Minervini, UCL  
Sebastião Miranda, Priberam Labs  
Preslav Nakov, Qatar Computing Research Institute, HBKU  
Tsuayoshi Okita, Kyushuu institute of technology university  
Girish Palshikar, Tata Consultancy Services Limited  
Michael Pust, Information Sciences Institute  
Brendan Roof, Allen Institute for Artificial Intelligence  
Dan Roth, University of Pennsylvania  
Masoud Rouhizadeh, Johns Hopkins University  
Irene Russo, ILC CNR  
Swapna Somasundaran, Educational Testing Services  
Oyvind Tafjord, AI2  
Niket Tandon, Allen Institute for Artificial Intelligence  
Zhucheng Tu, University of Waterloo  
Chuan-Ju Wang, Academia Sinica  
Lyndon White, The University of Western Australia  
Zhuoran Yu, Google  
Boliang Zhang, Rensselaer Polytechnic Institute  
Zhuosheng Zhang, Shanghai Jiao Tong University

## Table of Contents

<i>Abbreviation Explorer - an interactive system for pre-evaluation of Unsupervised Abbreviation Disambiguation</i>	
Manuel Ciosici and Ira Assent.....	1
<i>ADIDA: Automatic Dialect Identification for Arabic</i>	
Ossama Obeid, Mohammad Salameh, Houda Bouamor and Nizar Habash .....	6
<i>Enabling Search and Collaborative Assembly of Causal Interactions Extracted from Multilingual and Multi-domain Free Text</i>	
George C. G. Barbosa, Zechy Wong, Gus Hahn-Powell, Dane Bell, Rebecca Sharp, Marco A. Valenzuela-Escárcega and Mihai Surdeanu .....	12
<i>INS: An Interactive Chinese News Synthesis System</i>	
Hui Liu, Wentao Qin and Xiaojun Wan .....	18
<i>Learning to Respond to Mixed-code Queries using Bilingual Word Embeddings</i>	
Chia-Fang Ho, Jason Chang, Jih-Jie Chen and Chingyu Yang .....	24
<i>Train, Sort, Explain: Learning to Diagnose Translation Models</i>	
Robert Schwarzenberg, David Harbecke, Vivien Macketanz, Eleftherios Avramidis and Sebastian Möller.....	29
<i>compare-mt: A Tool for Holistic Comparison of Language Generation Systems</i>	
Graham Neubig, Zi-Yi Dou, Junjie Hu, Paul Michel, Danish Pruthi and Xinyi Wang .....	35
<i>Eidos, INDRA, &amp; Delphi: From Free Text to Executable Causal Models</i>	
Rebecca Sharp, Adarsh Pyarelal, Benjamin Gyori, Keith Alcock, Egoitz Laparra, Marco A. Valenzuela-Escárcega, Ajay Nagesh, Vikas Yadav, John Bachman, Zheng Tang, Heather Lent, Fan Luo, Mithun Paul, Steven Bethard, Kobus Barnard, Clayton Morrison and Mihai Surdeanu .....	42
<i>fairseq: A Fast, Extensible Toolkit for Sequence Modeling</i>	
Myle Ott, Sergey Edunov, Alexei Baevski, Angela Fan, Sam Gross, Nathan Ng, David Grangier and Michael Auli.....	48
<i>FLAIR: An Easy-to-Use Framework for State-of-the-Art NLP</i>	
Alan Akbik, Tanja Bergmann, Duncan Blythe, Kashif Rasul, Stefan Schweter and Roland Vollgraf	54
<i>ChatEval: A Tool for Chatbot Evaluation</i>	
Joao Sedoc, Daphne Ippolito, Arun Kirubarajan, Jai Thirani, Lyle Ungar and Chris Callison-Burch	60
<i>LeafNATS: An Open-Source Toolkit and Live Demo System for Neural Abstractive Text Summarization</i>	
Tian Shi, Ping Wang and Chandan K. Reddy .....	66
<i>End-to-End Open-Domain Question Answering with BERTserini</i>	
Wei Yang, Yuqing Xie, Aileen Lin, Xingyu Li, Luchen Tan, Kun Xiong, Ming Li and Jimmy Lin	72
<i>FAKTA: An Automatic End-to-End Fact Checking System</i>	
Moin Nadeem, Wei Fang, Brian Xu, Mitra Mohtarami and James Glass .....	78

<i>iComposer: An Automatic Songwriting System for Chinese Popular Music</i> Hsin-Pei Lee, Jhih-Sheng Fang and Wei-Yun Ma .....	84
<i>Plan, Write, and Revise: an Interactive System for Open-Domain Story Generation</i> Seraphina Goldfarb-Tarrant, Haining Feng and Nanyun Peng .....	89
<i>LT Expertfinder: An Evaluation Framework for Expert Finding Methods</i> Tim Fischer, Steffen Remus and Chris Biemann .....	98
<i>SkillBot: Towards Automatic Skill Development via User Demonstration</i> Yilin Shen, Avik Ray, Hongxia Jin and Sandeep Nama .....	105
<i>Multilingual Entity, Relation, Event and Human Value Extraction</i> Manling Li, Ying Lin, Joseph Hoover, Spencer Whitehead, Clare Voss, Morteza Dehghani and Heng Ji .....	110
<i>Litigation Analytics: Extracting and querying motions and orders from US federal courts</i> Thomas Vacek, Dezhao Song, Hugo Molina-Salgado, Ronald Teo, Conner Cowling and Frank Schilder .....	116
<i>Community lexical access for an endangered polysynthetic language: An electronic dictionary for St. Lawrence Island Yupik</i> Benjamin Hunt, Emily Chen, Sylvia L.R. Schreiner and Lane Schwartz .....	122
<i>Visualizing Inferred Morphotactic Systems</i> Haley Lepp, Olga Zamaraeva and Emily M. Bender .....	127
<i>A Research Platform for Multi-Robot Dialogue with Humans</i> Matthew Marge, Stephen Nogar, Cory J. Hayes, Stephanie M. Lukin, Jesse Bloecker, Eric Holder and Clare Voss .....	132
<i>Chat-crowd: A Dialog-based Platform for Visual Layout Composition</i> Paola Cascante-Bonilla, Xuwang Yin, Vicente Ordonez and Song Feng .....	138



# Conference Program

Monday, June 3, 2019

15:00–16:30 *Session I*

*Abbreviation Explorer - an interactive system for pre-evaluation of Unsupervised Abbreviation Disambiguation*

Manuel Ciosici and Ira Assent

*ADIDA: Automatic Dialect Identification for Arabic*

Ossama Obeid, Mohammad Salameh, Houda Bouamor and Nizar Habash

*Enabling Search and Collaborative Assembly of Causal Interactions Extracted from Multilingual and Multi-domain Free Text*

George C. G. Barbosa, Zechy Wong, Gus Hahn-Powell, Dane Bell, Rebecca Sharp, Marco A. Valenzuela-Escárcega and Mihai Surdeanu

*INS: An Interactive Chinese News Synthesis System*

Hui Liu, Wentao Qin and Xiaojun Wan

*Learning to Respond to Mixed-code Queries using Bilingual Word Embeddings*

Chia-Fang Ho, Jason Chang, Jih-Jie Chen and Chingyu Yang

*Train, Sort, Explain: Learning to Diagnose Translation Models*

Robert Schwarzenberg, David Harbecke, Vivien Macketanz, Eleftherios Avramidis and Sebastian Möller

17:00–18:30 *Session II*

*compare-mt: A Tool for Holistic Comparison of Language Generation Systems*

Graham Neubig, Zi-Yi Dou, Junjie Hu, Paul Michel, Danish Pruthi and Xinyi Wang

*Eidos, INDRA, & Delphi: From Free Text to Executable Causal Models*

Rebecca Sharp, Adarsh Pyarelal, Benjamin Gyori, Keith Alcock, Egoitz Laparra, Marco A. Valenzuela-Escárcega, Ajay Nagesh, Vikas Yadav, John Bachman, Zheng Tang, Heather Lent, Fan Luo, Mithun Paul, Steven Bethard, Kobus Barnard, Clayton Morrison and Mihai Surdeanu

*fairseq: A Fast, Extensible Toolkit for Sequence Modeling*

Myle Ott, Sergey Edunov, Alexei Baevski, Angela Fan, Sam Gross, Nathan Ng, David Grangier and Michael Auli

*FLAIR: An Easy-to-Use Framework for State-of-the-Art NLP*

Alan Akbik, Tanja Bergmann, Duncan Blythe, Kashif Rasul, Stefan Schweter and Roland Vollgraf

**Monday, June 3, 2019 (continued)**

*ChatEval: A Tool for Chatbot Evaluation*

Joao Sedoc, Daphne Ippolito, Arun Kirubarajan, Jai Thirani, Lyle Ungar and Chris Callison-Burch

*LeafNATS: An Open-Source Toolkit and Live Demo System for Neural Abstractive Text Summarization*

Tian Shi, Ping Wang and Chandan K. Reddy

**Tuesday, June 4, 2019**

**11:00–12:30** *Session III*

*End-to-End Open-Domain Question Answering with BERTserini*

Wei Yang, Yuqing Xie, Aileen Lin, Xingyu Li, Luchen Tan, Kun Xiong, Ming Li and Jimmy Lin

*FAKTA: An Automatic End-to-End Fact Checking System*

Moin Nadeem, Wei Fang, Brian Xu, Mitra Mohtarami and James Glass

*iComposer: An Automatic Songwriting System for Chinese Popular Music*

Hsin-Pei Lee, Jih-Sheng Fang and Wei-Yun Ma

*Plan, Write, and Revise: an Interactive System for Open-Domain Story Generation*

Seraphina Goldfarb-Tarrant, Haining Feng and Nanyun Peng

*LT Expertfinder: An Evaluation Framework for Expert Finding Methods*

Tim Fischer, Steffen Remus and Chris Biemann

*SkillBot: Towards Automatic Skill Development via User Demonstration*

Yilin Shen, Avik Ray, Hongxia Jin and Sandeep Nama

**15:30–17:00** *Session IV*

*Multilingual Entity, Relation, Event and Human Value Extraction*

Manling Li, Ying Lin, Joseph Hoover, Spencer Whitehead, Clare Voss, Morteza Dehghani and Heng Ji

*Litigation Analytics: Extracting and querying motions and orders from US federal courts*

Thomas Vacek, Dezhao Song, Hugo Molina-Salgado, Ronald Teo, Conner Cowling and Frank Schilder

**Tuesday, June 4, 2019 (continued)**

*Community lexical access for an endangered polysynthetic language: An electronic dictionary for St. Lawrence Island Yupik*

Benjamin Hunt, Emily Chen, Sylvia L.R. Schreiner and Lane Schwartz

*Visualizing Inferred Morphotactic Systems*

Haley Lepp, Olga Zamaraeva and Emily M. Bender

*A Research Platform for Multi-Robot Dialogue with Humans*

Matthew Marge, Stephen Nogar, Cory J. Hayes, Stephanie M. Lukin, Jesse Bloecker, Eric Holder and Clare Voss

*Chat-crowd: A Dialog-based Platform for Visual Layout Composition*

Paola Cascante-Bonilla, Xuwang Yin, Vicente Ordonez and Song Feng

